University of Ibadan
Journal of Science and Logics in ICT Research

Crop Recommendation Analysis and Validation in Nigeria Using Machine Learning Algorithms

[1]Samuel, K. S., [2]Olamiti, A. O. and [3]Ayorinde, I. T.

[1,2&3]Department of Computer Science, University of Ibadan, Ibadan, Nigeria.
Emails: [1]kavalgasam@gmail.com ; [2]aolamiti@yahoo.com ; [3]temiayorinde@yahoo.com

Abstract

Crop recommendation systems are crucial for optimizing agriculture by suggesting crops based on environmental and soil conditions. Failure in selecting suitable crops can result in low yields and resource wastage. This study builds an improved recommendation system for Nigerian farmers. Data from various sources, including the Nigeria Metrological Agency, the Agronomy Department University of Ibadan, Ahmadu Bello University Zaria, and Federal University Wukari, were preprocessed using numpy and pandas. The climate parameters used were Rainfall, Temperature and Humidity while the soil parameters were Nitrogen (N), Phosphorus (P), Potassium (K), Calcium (C) and Magnesium (Mg). The pH was used to measure the soil acidity or alkalinity. The 18 crops considered were Bambara Nut, Cassava, Cocoyam, Tomato, Yam, Acha, Cocoa, Beans, Groundnut, Beniseed, Maiza, Rice, Oil Palm, Cashew, Sugar cane, Sweet Potato, Pepper and Coconut. After preprocessing, the dataset was partitioned into training, validation, and testing sets in the ratio 80:10:10. Four Machine learning algorithms which are Random Forest, Naïve Bayes, K Nearest Neighbor, and Support Vector Machine (SVM) were employed, with Random Forest outperforming others in accuracy, precision, recall, and F1 score. Naïve Bayes ranked second, followed by K Nearest Neighbor, and Support Vector Machine performed as the poorest. The models effectively recommended crops for specific climates and soils, with SVM being the least effective. Hence, this study demonstrates the importance of accurate crop recommendations in maximizing agricultural productivity.

Keywords: *Climate data, Soil nutrients, Machine learning, Crop selection predictive model, Agricultural productivity.*

1. INTRODUCTION

Agriculture holds utmost significance as our planet's primary sector for sustenance while simultaneously playing a pivotal role in providing raw materials for other industrial endeavours. Agricultural sector employs a large portion of the work force, and contributes significantly to the Gross Domestic Product (GDP), making it an important backbone of Nigeria's economy. The sector is being confronted with challenges from antiquated practices, soil quality, limited funding, inadequate infrastructure, and the adverse effects of climate change. These challenges have led to the inability of the agricultural production to meet the demands of the ever-expanding population, potentially leading to a deficit in the global food supply. Also, the augmentation of food production is paramount for emerging nations with limited land and resources. Hence, it becomes imperative to meticulously choose an appropriate crop for a specific region to augment its production rate [1].

According to Ayorinde [2], crops are plants' products grown to provide food, fuel and clothing among others. Agriculture requires farmers to cultivate profitable and sustainable crops. Not cultivating the right crop in the right environment can significantly impact crop yield, leading to decreased productivity and potential financial losses for farmers [3]. The agricultural sector contributes 10–20% of the total anthropogenic greenhouse gas (GHG) emissions. Consequently, climate change can negatively affect crop yields and livestock production, thus threatening food security, especially on vulnerable continents like Africa [4]. When farmers overlook essential factors like climate appropriateness, soil quality, and

Samuel, K. S., Olamiti, A. O. and Ayorinde, I. T. (2024). Crop Recommendation Analysis and Validation in Nigeria Using Machine Learning Algorithms. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 11 No. 2, pp. 36 – 45.

market demand, the selected crops might face difficulties in flourishing and achieving their highest possible yield. Inappropriate crop choices could lead to challenges in adjusting to the local climate, leading to stunted growth, heightened susceptibility to pests and diseases, and ultimately diminished yields [3].

Climate change refers to the increase in greenhouse gas emissions (GHG) such as nitrous oxide (N_2O), carbon dioxide (CO_2), and methane (CH_4) in the atmosphere, causing irregularity, variability, and unpredictability of rainfall, temperature increase, floods, and drought [4], which can drastically affect farming activities.

Soil is a critical part of thriving agriculture and the source of nutrients used to grow crops. The productive capacity of soil is dependent on soil fertility. Soil fertility is the ability of soil to supply essential plant nutrients and soil water in adequate amounts and proportions for plant growth and reproduction without toxic substances that may inhibit plant growth [5].

Adequate knowledge about climate change and soil properties is important to help farmers use machine learning and decide what crops to grow, how to care for those crops throughout the growing season, and how to predict crop yields [6]. Crop recommendation in agriculture involves the use of data-driven technologies and algorithms to suggest the most suitable crops for cultivation in a specific region. This process takes into account various factors such as Agro-climatic conditions, soil properties, historical crop performance, and other relevant parameters. The goal is to provide farmers with tailored guidance, helping them make informed decisions about what to plant to optimize yields and resource utilization.

2. Related Works

In investigating crop recommendation and validation systems, knowledge from previous scholarly endeavours investigating the complexities of crop recommendation analysis by utilizing machine learning algorithms, intelligent farming techniques, and Artificial Intelligence (AI) in the agricultural sector is considered. These investigations offer valuable groundwork for comprehending the obstacles, methodologies, and significance of crop recommendation in agriculture.

Apat *et. al.* [7] introduced a novel AI-based recommendation system for precision agriculture, leveraging machine learning algorithms and optimization techniques. This system stands out for its high accuracy in crop recommendation, a feat achieved by incorporating key features such as nitrogen, potassium, phosphorus, temperature, humidity, pH, and rainfall. They applied data balancing techniques like SMOTE to enhance the system's performance. However, the authors suggested that future research should be focused on improving the accuracy of the model by incorporating additional environmental factors and using more advanced machine learning algorithms. Future research could also focus on improving the accuracy of the models, developing a user-friendly interface, chatbots, and other digital tools to help farmers make informed decisions about crop selection.

Patel *et. al.* [8] developed and applied an intelligent system that can suggest suitable crops for farmers in India; their aim is to address and help farmers pick the best crops for their situation and environment by predicting which crops fit well with the factors that influence crop growth, such as soil nutrients like nitrogen, phosphorous, potassium, pH, and climate factors including temperature, humidity and rainfall, employing Decision Tree, Support Vector Machine (SVM), Logistic Regression (LR), and GaussianNB. They achieved a training accuracy of 99.5% and a validation accuracy of 99.3%. Their work can be enhanced by employing a more significant number of attributes in the dataset and by building a model that can classify healthy and diseased crop leaves. If the crop has any disease, predict which disease it is.

Pachade and Sharma [9] focused their research on machine learning for weather-specific crop recommendations. Their system recommends crops for farmers to grow based on input from the field provided by the farmer, such as temperature, soil, moisture, and nutrients like nitrogen, phosphorus, potassium, pH, and rainfall. Three popular machine learning algorithms were used in this study: decision tree, logistics regression, and random forest. The Random Forest among them demonstrated the highest outcomes with 99.32% accuracy. While this result is encouraging, the authors acknowledged that the system needs all of the data collected from crops to make the system even better and assist farmers in making conclusions about which crops to cultivate.

Dey *et. al.* [10] employed machine learning models to produce actionable suggestions regarding crop selection and nutrient assessment, leveraging factors such as nitrogen, phosphorus, potassium, soil acidity (pH), and climatic conditions in India. The dataset used in their research consists of information on nitrogen, potassium, phosphorus, soil pH, and climatic factors for eleven agricultural and ten horticultural crops. The models were trained using this dataset to provide accurate crop selection and nutrient determination recommendations. The research assessed the effectiveness of five machine learning models (Support Vector Machine, XGBoost, Random Forest, K-Nearest Neighbors, and Decision Tree) in providing practical recommendations for crop selection or determining necessary nutrients in a specific location. The XGBoost model exhibited the highest accuracy, achieving precision rates of 99.09% for crop recommendations, 99.3% for horticultural crops, and 98.51% for a combined set. The study suggests that individual datasets should be evaluated separately for each crop category (agricultural and horticultural) rather than merging datasets to enhance predictive accuracy.

Priya *et. al.* [11] proposed a crop recommendation system based on the SVM algorithm to address the issue of farmers' inappropriate crop selection, which leads to reduced productivity and food shortages. The system analyzes soil parameters and predicts the most suitable crop, helping farmers eliminate losses and increase productivity. The accuracy of the proposed algorithm is reported to be 97%, indicating its effectiveness in crop classification and recommendation. The paper highlights the importance of considering soil N, P, K, and pH values in determining the most productive crops for specific conditions. The system incorporates data visualization techniques to understand the relationships between different variables and identify patterns in the data. The proposed system can be further enhanced by integrating Internet of Things (IoT) technologies to obtain real-time soil values, enabling more accurate and timely crop recommendations.

Rani *et. al.* [12] proposes a machine learning-based crop selection model that considers climate conditions and soil parameters for optimal crop selection in intelligent agriculture. The authors highlight the importance of analyzing agro-climatic conditions to determine the right crop for a specific region and season. The model they used consists of two phases: weather prediction using Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) and crop selection using a Random Forest Classifier. They emphasize the superior performance of their model compared to traditional methods, with high accuracy in weather prediction and crop selection. LSTM RNN is employed for weather prediction. A Random Forest Classifier is used for crop selection, taking into account multiple weather and soil parameters. It accurately identifies the optimal crop and predicts resource dependency and sowing time. LSTM RNN showed better weather prediction results than traditional methods like ANN. For future work, the authors recommended conducting experiments and validation in different geographical areas to assess the generalization and robustness of the model.

Ali [13] proposed a system aimed at aiding farmers in Pakistan to maximize crop yields. It suggests employing Machine Learning, particularly linear regression analysis, to develop a crop recommendation system tailored for local farmers. This system utilizes meteorological data and agricultural information from diverse sources, including government departments and research institutions, to analyze and predict suitable crops based on temperature, thereby enhancing crop production. While the dataset used is specific to the Nawabshah region in Sindh, Pakistan, the methodology is adaptable for other areas. The system primarily focuses on suggesting crops based on climatic conditions and soil characteristics to boost yield for farmers. They achieved an average accuracy of 90% on the dataset provided, surpassing existing systems in accuracy. However, the system has limitations and potential areas for future development. Expanding the dataset beyond Nawabshah to include data from other regions in Pakistan would broaden the system's applicability.

3. Methodology

Figure 1 shows the overall design of the crop recommendation system. The dataset for this study is secondary data and was collected from Nigeria Metrological Agency (NIMET); Agronomy Department, University of Ibadan; Department of Agriculture, Ahmadu Bello University Zaria; and Department of Agriculture, Federal University Wukari. They were preprocessed using numpy and pandas in Jupyter notebook and the preprocessed dataset was partitioned into training, validation and testing sets in the ratio 80:10:10. The machine learning algorithms used were Random Forest, Naïve Bayes, K Nearest Neighbor, and Support Vector Machine. The performance metrics used for the models were Accuracy, Precision, Recall and F1 score. The dataset contains the following features: Rainfall, Temperature, Humidity, Soil pH, Phosphorus (K), potassium (P), Nitrogen (N), Calcium (Ca), Magnesium (Mg), and crop types (label), totalling 4500 rows and ten features. It encompasses 18 different crops: Bambara nut, Groundnut, Maize, Rice, Beniseed, Yam, Acha, Cassava, Oil palm, Cashew, Sugarcane, Sweet potato, Pepper, Cocoyam, Tomato, Beans, Cocoa and Coconut.

3.1 Preprocessing

To successfully implement these algorithms, preprocessing is required to get the relevant features. The dataset, collected from different sources, is sometimes in raw form; it is sometimes incomplete, inconsistent, and contains reductant data. Therefore, redundant data should be filtered, normalized, or standardized before developing machine learning models.

3.2 Checking for missing values in the dataset.

During preprocessing, missing or incomplete entries are taking care of. Finding missing values is essential because they might compromise the precision and dependability of modelling and data analysis. The missing data in the dataset can be handled in several ways, including removing partial information, imputation using statistical measures like mean or median, or more sophisticated approaches like predictive modelling. In the end, partial or incomplete information were removed to preserve the dataset's integrity and improve the model's performance.

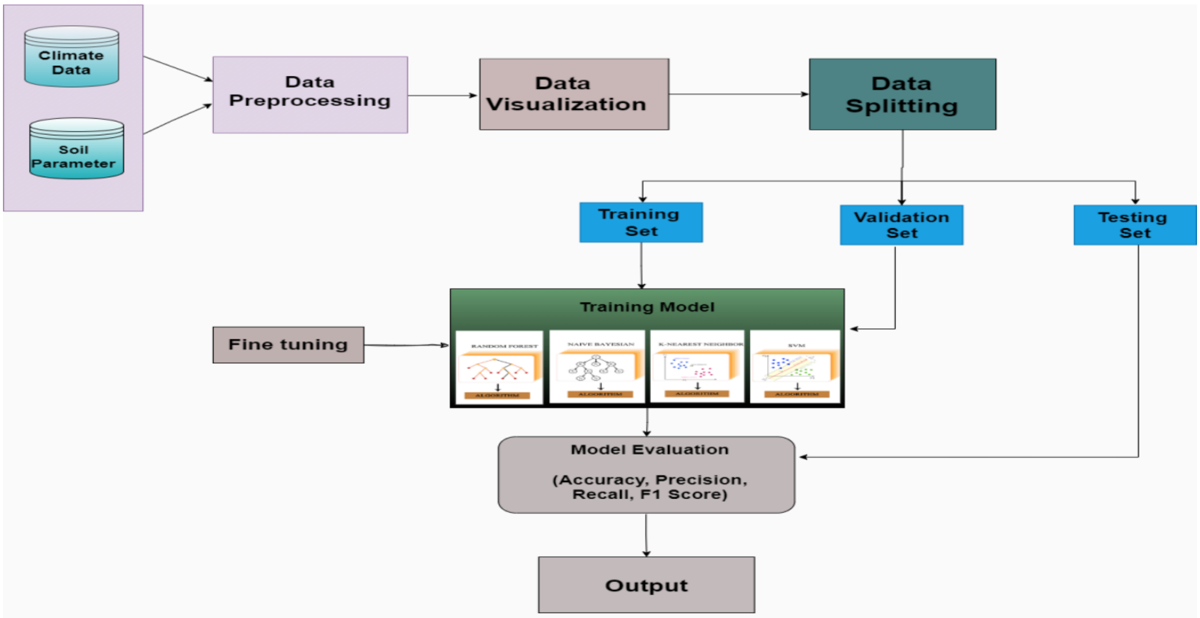

Figure 1. Diagram of The Crop Recommendation Analysis and Validation System

1.3 Converting categorical data to numerical values

In this study, the dataset comprises of categorical data, necessitating the conversion of these categories into numerical values using a label encoder. This technique assigns a distinct integer to each category. For example, the first crop on our label, Bambara nut, is assigned a unique integer 0, and groundnut is assigned a unique integer 1 until all the crops are assigned unique values.

3.4 Splitting the dataset into training, validation, and testing set

To facilitate the training of machine learning models, it is necessary to divide the dataset into separate training, validation and testing sets in the ratio 80:10:10. This division aims to ensure that the model can effectively apply the knowledge gained from the training data to new and previously unseen examples. By evaluating the model's performance on untested data, we can determine its ability to generalize beyond the examples it was trained on.

4. Implementation

This section discusses how the four machine learning algorithms were implemented.

4.1 Random Forest Model creation

The required libraries (like sci-kit-learn in Python) were imported and a random forest classifier was created. Other steps taken were:

i. **Initialize Random Forest**: This creates an instance of the RandomForestClassifier class with default hyperparameters.
ii. **Training**: During this stage, the classifier undergoes training utilizing the training data, where X_train represents the features and Y_train denotes the labels.
iii. **Define hyperparameter**: This creates a dictionary specifying the hyperparameters to be tuned (n_estimators and max_depth) and the range of values to try for each hyperparameter. The grid search will perform an exhaustive search over a specified parameter grid to find the best parameters for a model. It uses cross-validation (cv=5 indicates 5-fold cross-validation) to evaluate each combination of hyperparameters and choose the best combination of values.
iv. **Model Training**: This fits the random forest model to the training data using the fit method. Create a random bootstrap sample from the training data (with replacement) for each tree. Build a decision tree on the bootstrap sample, considering the subset of features at each split (random feature selection).
v. **Ensemble Creation**: This combines the predictions of all decision trees through majority voting (classification).
vi. **Prediction**: The ensemble of the decision trees were used to make predictions on new, unseen data.
vii. **Output**: The final Random Forest model predicts the given task.
viii. **Model Evaluation**: The model's performance was evaluated on the test data using appropriate metrics (accuracy, precision, recall, and F1 score for classification).
ix. **Predictions**: Once the model is trained, it is used to predict new, unseen data.

4.2 Naive Bayes Model creation

i. **Initialization**: This creates an instance of the GaussianNB class, which implements the Gaussian Naive Bayes algorithm for classification.
ii. **Training**: The Gaussian Naive Bayes classifier was trained using the training data (X_train for features and Y_train for labels).
iii. **Make Predictions on Test Data**: Using the trained classifier; predictions were made on the test data (X_test). The prediction method takes the features of the test set as input and returns the predicted labels.
iv. **Model Evaluation**: This calculates the accuracy of the Gaussian Naive Bayes classifier's predictions by comparing the predicted label to the actual labels using the accuracy_score function. The accuracy_score function calculates the accuracy as the fraction of correctly classified sample sets that were classified correctly out of the total number of samples.
v. **Continuous Improvement**: To improve the model's relevance and accuracy, new data were added to it regularly

4.3 Support Vector Machine Model creation

i. **Initialization**: Initialize a support vector classifier without specifying any hyperparameters.

ii. **Training**: Trains the classifier on the training data X_train with corresponding labels Y_train.

iii. **Hyperparameter tuning**: Svc-params define a dictionary svc-params containing the hyperparameter to be tuned. It specifies different values for the regularization parameter and the kernel type (linear and rbf). Set up a grid search cross-validation to find the best combination of hyperparameters for the SVM classifier.

iv. **Prediction**: For a new location with unknown suitable crops - Input the location's features into the trained SVM model. The model predicts the crop category (or multiple categories) with the highest probability of success based on the learned hyperplanes.

v. **Recommendation and Refinement**: The system recommends crops based on the SVM's prediction, potentially considering additional factors or user preferences. As farmers adopt the recommendations and provide feedback on their experiences, the system can be refined by incorporating new data to improve future predictions.

4.4. K-Nearest Neighbor Model creation

i. **Initialize K-Nearest Neighbors Classifier**: creates an instance of the K Neighbors Classifier class with default hyperparameters. K Neighbors Classifier is a classification algorithm that works by finding the 'k' nearest data points in the feature space and making predictions based on the labels of those points.

ii. **Training**: Train the KNN classifier using the training data (X_train for features and Y_train for labels).

iii. **Define Hyperparameter Grid**: Here, a dictionary, knn_params, specifies the hyperparameters to be tuned (n_neighbors) and the range of values to try for the number of neighbors. GridSearchCV is used again with the KNN classifier and the defined hyperparameter grid. Grid search searched through the parameter grid and find the best combination of hyperparameters based on the model's performance on the training data.

iv. **Prediction**: Predictions are generated on the test data (X_test) with the optimal estimator identified.

4.5 Evaluate Model Accuracy

The model's accuracy is evaluated by comparing the predicted labels (knn_pred) to the actual labels (Y_test) using the accuracy_score function.

4.6 Machine Learning Model Performance Metrics

Evaluating the effectiveness and suitability of the machine learning (ML) algorithms is crucial for various tasks. This process involves assessing how well a model performs on specific datasets, utilizing metrics such as accuracy, precision, recall, and F1-score to gauge their performance.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions\ Made}$$

Accuracy indicates the proportion of correct predictions made by the model, providing a straightforward measure that is easy to comprehend and interpret. However, it can be misleading in an imbalanced dataset because, in real-world agriculture, some crops are more common than others. In such cases, focusing solely on accuracy can be misleading; we will also use precision, recall, and F1 scores to evaluate how well the models performed.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Precision would measure the system's accuracy in recommending crops in the three zones. High precision will indicate that the recommended crops are more likely suitable for the given conditions.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

The recall will ensure that the system identifies and recommends all the suitable crops for a particular region and captures all relevant crops, minimizing false negatives.

$$F1\ Score = \frac{2\ X\ Precison\ X\ Recall}{Precision + Recall}$$

The F1-Score is utilized when there is a need to balance precision and recall. It furnishes a metric that takes into account both false positives and negatives, resulting in a more holistic assessment.

5.0 Discussion of Results

The various results are discussed in this section

5.1 Results of the Standard Metrics

Table 1 shows that Random Forest Model, Naïve Bayes, K Nearest Neighbor and Support Vector Machine gave an accuracy of 99.9, 99.8, 96.1 and 78.1%, respectively; Precision of 97.0, 96.0, 76.0 and 48%, respectively; Recall of 99.0, 97.0, 79.0 and 48.0%; respectively and F1 Score of 99.0, 98.0, 86.0 and 52.0%, respectively. These indicate that Random Forest was best in recommending appropriate crop for the right soil and climate, followed by Naïve Bayes, KNN with SVM as the least because the higher the values of accuracy, precision, recall and F1 score, the better the model performance.

5.2 Scatter plots

Scatter plot is a graphical representation used to display the relationship between two variables and identify patterns or trends in the data. It shows the performance of the machine learning models on the crop recommendation task as shown in Figures 2 to 5. The x-axis on the scatter plot is labelled "Actual crop name" and the y-axis is labelled "Predicted crop name". the diagonal line from the bottom left to the top right of the graph represents perfect performance where the model would correctly predict the crop name for every sample.

The data point in the graph represents the number of times that the model predicted a particular crop name. The closer a data point is to the diagonal line, the better the model performed for that crop, data point far from the diagonal line shows misclassification. In considering the scatter plot in Figures 2 and 3, Random Forest and Naive Bayes models appear to be very effective at classifying crops with the dataset.

In Figure 4, the scatter plot suggest that K Nearest Neighbor model performed fairly at predicting some crops' names, because there are many data points that cluster around the diagonal line, and there are also data points far away from the diagonal line, which suggests that the model prediction was less accurate compared to Random Forest and Naïve Bayes.

In Figure 5, the scatter plot suggests that most of the data points fall below the diagonal line, which indicates that the model with SVM has weaker performance and difficulty in correctly classifying some crops.

Table 1: Results of All the Models.

S/N	**MODELS**	**ACCURACY (%)**	**PRECISION (%)**	**RECALL (%)**	**F1 SCORE (%)**
1	RANDOM FOREST	99.85	97	99	99
2	NAÏVE BAYES	99.77	96	97	98
3	K NEAREST NEIGHBOR	96.14	79	79	82
4	SUPPORT VECTOR MACHINE	78.07	43	48	52

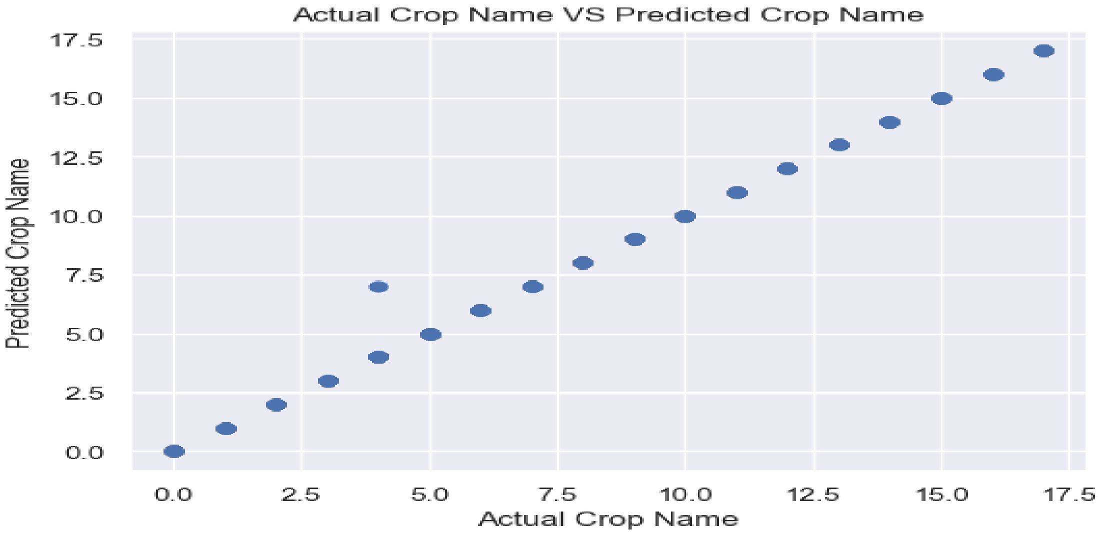

Figure 2: Diagram of a Scatter Plot Showing Performance for Random Forest Model.

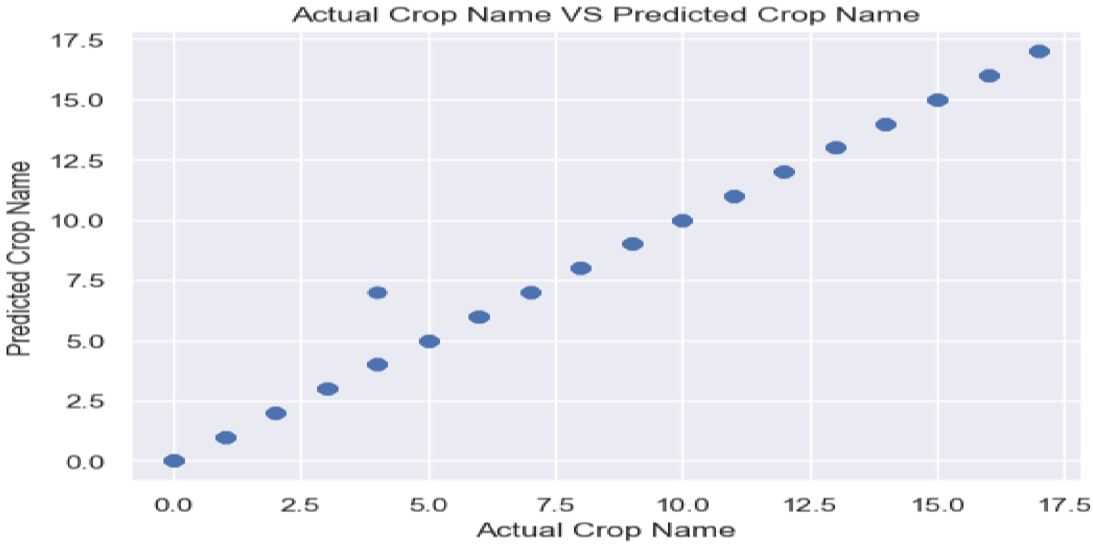

Figure 3: Diagram of a Scatter Plot Showing Performance for Naïve Bayes Model.

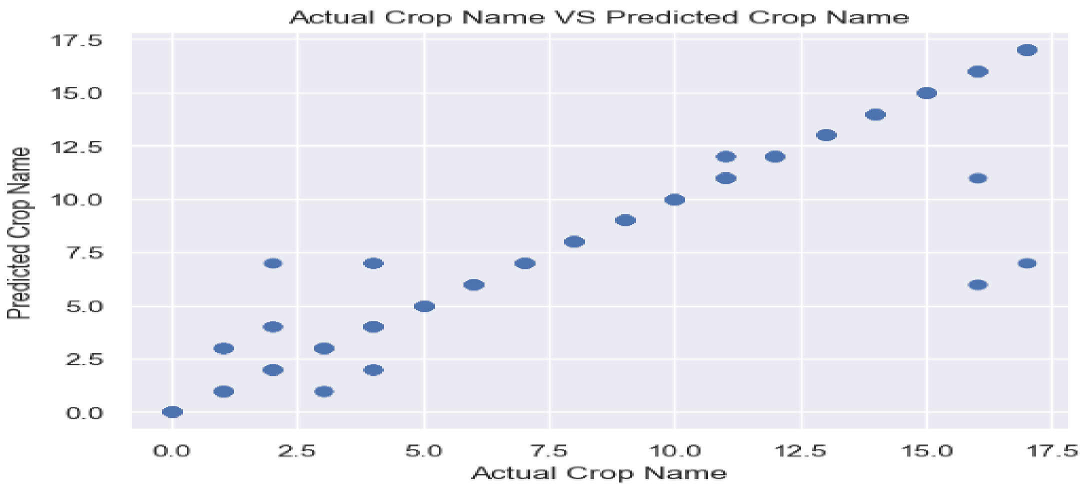

Figure 4: Diagram of a Scatter Plot Showing Performance of KNN model.

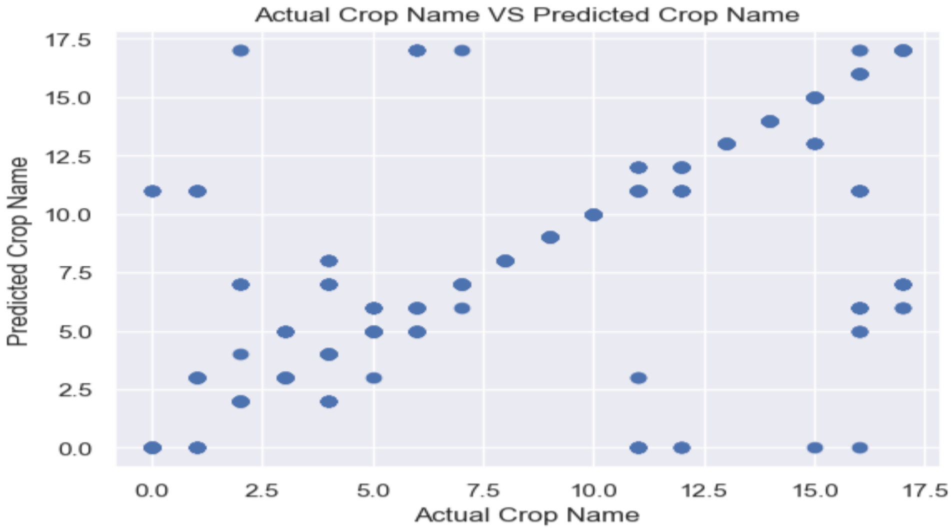

Figure 5: Diagram of a scatter plot showing performance of SVM models.

6. Conclusion

Random Forest performs exceptionally well across all metrics, demonstrating high accuracy, precision, recall, and F1 score. Naive Bayes also performs well across all metrics, demonstrating high accuracy, precision, recall, and F1 score. KNN performs reasonably well, with good accuracy and F1 score, but slightly lower precision and recall than Random Forest and Naive Bayes. SVM performs relatively poorly compared to the other models, with lower accuracy, precision, recall, and F1 score. It may not be the best choice for this specific task or require further optimization.

Based on the provided metrics, Random Forest and Naive Bayes demonstrate superior performance, followed by KNN, while SVM shows relatively weaker performance. Hence, from this study, Random Forest and Naive Bayes are robust and can be deployed to any region while K Nearest Neighbor and Support Vector Machine might be best suited for region specific dataset because they performed well in existing models with region specific datasets. This conclusion may guide the selection of the most suitable model for the crop recommendation analysis and validation system.

References

[1] Hasan, M., Marjan, A., & Uddin, P. (2023). Ensemble machine learning- based recommendation system for effective prediction of suitable agricultural crop cultivation. August, 1–18. https://doi.org/10.3389/fpls.2023.1234555

[2] Ayorinde, I. T. (2023). Visualization of a Domain Knowledge (A Case Study of Crops Ontology). Knowledge Graph and Semantic Web Conference (2023): Joint Proceedings of Second International Workshop on Semantic Reasoning and Representation in IoT (SWIoT-2023) and Third International Workshop on Multilingual Semantic Web (MSW-2023). University of Zaragoza, Zaragoza, Spain. November 13-15, 2023.

[3] Dahiphale, D., Shinde, P., Patil, K., & Dahiphale, V. (2023). Smart Farming: Crop Recommendation using Machine Learning with Challenges and Future Ideas. https://doi.org/10.36227/techrxiv.23504496.v1

[4] Obaji, A., Dimon, E., Nambeye, E., Diouf, N., Adelusi, O., & Boudalia, S. (2022). Climate-Smart Agriculture in African Countries: A Review of Strategies and Impacts on Smallholder Farmers. Sustainability, 14, 11370. https://doi.org/10.3390/su141811370

[5] Raut, J., & Mittal, S. (2020). Soil Fertility and Crop Recommendation using Machine Learning and Deep Learning Techniques: A Review. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 11, 1119–1127

[6] Lahza, H., Naveen Kumar, K. R., Sreenivasa, B. R., Shawly, T., Alsheikhy, A. A., Hiremath, A. K., & Lahza, H. F. M. (2023). Optimization of Crop Recommendations Using Novel Machine Learning Techniques. Sustainability (Switzerland), 15(11). https://doi.org/10.3390/su15118836

[7] Apat, S. K., Mishra, J., Raju, K. S., & Padhy, N. (2023). An Artificial Intelligence-based Crop

Recommendation System using Machine Learning. Journal of Scientific and Industrial Research, 82(5), 558–567. https://doi.org/10.56042/jsir.v82i05.1092

[8] Patel, M., Rane, A., & Patni, V. (2023.). Crop Recommendation System using Machine Learning. https://www.kaggle.com/datasets/atharvaingle/crop-

[9] Pachade, R., & Sharma, A. (2022). Machine learning for weather-specific crop recommendation. International Journal of Health Sciences, 4527–4537. https://doi.org/10.53730/ijhs.v6nS8.13222

[10] Dey, B., Ferdous, J., & Ahmed, R. (2024). Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables. Heliyon, 10(3), e25112. https://doi.org/10.1016/j.heliyon.2024.e25112

[11] Priya, V. S., Ramana, B. V., Kumar, A. S., Tanvireddy, A., & Jasta, A. (2023). Crop Recommendation Using Machine Learning. 2023 IEEE International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering, RMKMATE 2023, 2–6. https://doi.org/10.1109/RMKMATE59243.2023.10369475

[12] Rani, S., Mishra, A. K., Kataria, A., Mallik, S., & Qin, H. (2023). Machine learning-based optimal crop selection system in smart agriculture. Scientific Reports, 13(1), 1–11. https://doi.org/10.1038/s41598-023-42356-

[13] Ali, S. M. (2021). Machine Learning based Crop Recommendation System for Local Farmers of Pakistan. Revista Gestão Inovação e Tecnologias, November, 5735–5746. https://doi.org/10.47059/revistageintec.v11i4.2613